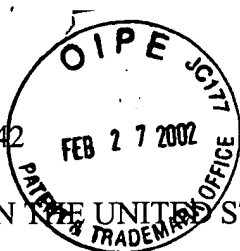


00862.022442



PATENT APPLICATION

#4  
03 262

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

AKIHIRO KUSHIDA, ET AL.

Application No.: 09/993,570

Filed: November 27, 2001

For: SPEECH RECOGNITION SYSTEM,  
SPEECH RECOGNITION CLIENT,  
THEIR CONTROL METHOD, AND  
COMPUTER READABLE MEMORY

)  
:  
Examiner: Unassigned  
)  
:  
Group Art Unit: 2641  
)  
:  
)  
:  
February 27, 2002  
)  
:  
)  
:  
:

Commissioner for Patents  
Washington, D.C. 20231

**RECEIVED**  
MAR '01 2002  
**Technology Center 2600**

SUBMISSION OF PRIORITY DOCUMENT

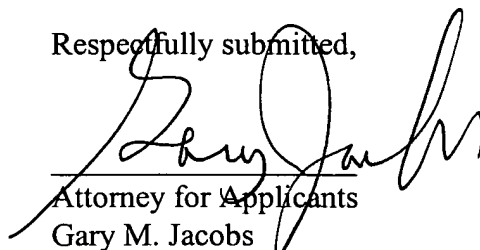
Sir:

In support of Applicants claim for priority under 35 U.S.C. § 119, enclosed is a  
certified copy of the following foreign application:

Japanese Patent Application No. 2000-360203, filed November 27, 2000.

Applicants' undersigned attorney may be reached in our Washington, D.C. office  
by telephone at (202) 530-1010. All correspondence should continue to be directed to our  
address given below.

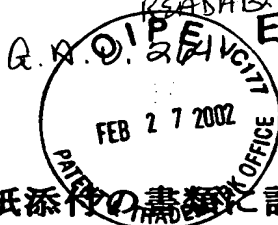
Respectfully submitted,

  
\_\_\_\_\_  
Attorney for Applicants  
Gary M. Jacobs  
Registration No. 28,861

FITZPATRICK, CELLA, HARPER & SCINTO  
30 Rockefeller Plaza  
New York, New York 10112-3801  
Facsimile: (212) 218-2200  
DC\_MAIN 88720v1

091993,57D  
AKIHIRO KUSHIDA, ET AL.  
"SPEECH RECOGNITION SYSTEM,  
SPEECH RECOGNITION CLIENT, THEIR  
CONTROL METHOD, AND COMPUTER  
READABLE MEMORY"

CFM 2442 US



日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office

出願年月日

Date of Application:

2000年11月27日

出願番号

Application Number:

特願2000-360203

出願人

Applicant(s):

キヤノン株式会社

RECEIVED

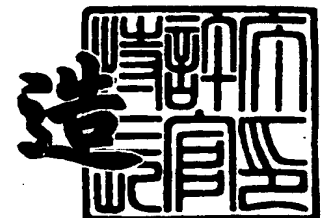
MAR 01 2002

Technology Center 2600

2001年12月21日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



出証番号 出証特2001-3110555

【書類名】 特許願

【整理番号】 4203021

【提出日】 平成12年11月27日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 1/00

【発明の名称】 音声認識システム及びその制御方法、コンピュータ可読メモリ

【請求項の数】 15

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 櫛田 晃弘

【発明者】

    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社  
社内

    【氏名】 小坂 哲夫

【特許出願人】

    【識別番号】 000001007

    【氏名又は名称】 キヤノン株式会社

【代理人】

    【識別番号】 100076428

    【弁理士】

    【氏名又は名称】 大塚 康德

    【電話番号】 03-5276-3241

【選任した代理人】

    【識別番号】 100101306

    【弁理士】

    【氏名又は名称】 丸山 幸雄

    【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100115071

【弁理士】

【氏名又は名称】 大塚 康弘

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0001010

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声認識システム及びその制御方法、コンピュータ可読メモリ

【特許請求の範囲】

【請求項 1】 クライアントで入力された音声をサーバで認識するクライアント・サーバ型の音声認識システムであって、

前記クライアントは、

音声を入力する音声入力手段と、

ユーザによって指定された認識対象語彙を登録することによって構成されるユーザ辞書を保持するユーザ辞書保持手段と、

前記音声入力手段で入力された音声データと、前記音声データの認識に使用する認識辞書の認識分野を決定するための辞書管理情報と、前記ユーザ辞書を前記サーバへ送信する送信手段と、

前記サーバは、

認識分野別に用意された複数種類の認識辞書群を保持する認識辞書保持手段と、

前記クライアントから受信した辞書管理情報に対応する認識辞書を前記複数種類の認識辞書群から決定する決定手段と、

前記決定手段で決定された認識辞書を少なくとも使用して、前記音声データの認識を行う認識手段と

を備えることを特徴とする音声認識システム。

【請求項 2】 前記認識手段は、前記決定手段で決定された認識辞書と前記クライアントから受信したユーザ辞書を使用して、前記音声データを認識することを特徴とする請求項 1 に記載の音声認識システム。

【請求項 3】 前記音声入力手段は、音声入力対象の入力フォームを表示する表示手段を備え、

前記辞書管理情報は、前記入力フォームの種類を示す入力フォーム識別子である

ことを特徴とする請求項 1 に記載の音声認識システム。

【請求項 4】 前記辞書管理情報は、更に、前記音声データの認識において

前記ユーザ辞書の使用の有無を示す情報を含む

ことを特徴とする請求項 1 に記載の音声認識システム。

【請求項 5】 前記ユーザ辞書における認識対象語彙は、認識対象語彙の読みと表記とが対応づけられて構成される

ことを特徴とする請求項 1 に記載の音声認識システム。

【請求項 6】 前記ユーザ辞書は、更に、前記入力フォーム識別子の少なくとも 1 つと、前記認識対象語彙とが対応づけられて構成される

ことを特徴とする請求項 3 に記載の音声認識システム。

【請求項 7】 前記ユーザ辞書は、更に、前記複数種類の認識辞書群のそれぞれの認識分野を示す認識辞書識別子の少なくとも 1 つと、前記認識対象語彙とが対応づけられて構成される

ことを特徴とする請求項 1 に記載の音声認識システム。

【請求項 8】 クライアントで入力された音声サーバで認識するクライアント・サーバ型の音声認識システムの制御方法であって、

音声を入力する音声入力工程と、

ユーザによって指定された認識対象語彙を登録することによって構成されるユーザ辞書を前記クライアントで保持するユーザ辞書保持工程と、

前記音声入力工程で入力された音声データと、前記音声データの認識に使用する認識辞書の認識分野を決定するための辞書管理情報と、前記ユーザ辞書を前記サーバへ送信する送信工程と、

認識分野別に用意された複数種類の認識辞書群を前記サーバで保持する認識辞書保持工程と、

前記クライアントから受信した辞書管理情報に対応する認識辞書を前記複数種類の認識辞書群から決定する決定工程と、

前記決定工程で決定された認識辞書を少なくとも使用して、前記音声データの認識を行う認識工程と

を備えることを特徴とする音声認識システムの制御方法。

【請求項 9】 前記認識工程は、前記決定工程で決定された認識辞書と前記クライアントから受信したユーザ辞書を使用して、前記音声データを認識する

ことを特徴とする請求項8に記載の音声認識システムの制御方法。

【請求項10】 前記音声入力工程は、音声入力対象の入力フォームを表示する表示工程とを備え、

前記辞書管理情報は、前記入力フォームの種類を示す入力フォーム識別子である

ことを特徴とする請求項8に記載の音声認識システムの制御方法。

【請求項11】 前記辞書管理情報は、更に、前記音声データの認識において前記ユーザ辞書の使用の有無を示す情報を含む

ことを特徴とする請求項8に記載の音声認識システムの制御方法。

【請求項12】 前記ユーザ辞書における認識対象語彙は、認識対象語彙の読みと表記とが対応づけられて構成される

ことを特徴とする請求項8に記載の音声認識システムの制御方法。

【請求項13】 前記ユーザ辞書は、更に、前記入力フォーム識別子の少なくとも1つと、前記認識対象語彙とが対応づけられて構成される

ことを特徴とする請求項10に記載の音声認識システムの制御方法。

【請求項14】 前記ユーザ辞書は、更に、前記複数種類の認識辞書群のそれぞれの認識分野を示す認識辞書識別子の少なくとも1つと、前記認識対象語彙とが対応づけられて構成される

ことを特徴とする請求項8に記載の音声認識システムの制御方法。

【請求項15】 クライアントで入力された音声サーバで認識するクライアント・サーバ型の音声認識システムの制御のプログラムコードが格納されたコンピュータ可読メモリであって、

音声を入力する音声入力工程と、

ユーザによって指定された認識対象語彙を登録することによって構成されるユーザ辞書を前記クライアントで保持するユーザ辞書保持工程のプログラムコードと、

前記音声入力工程で入力された音声データと、前記音声データの認識に使用する認識辞書の認識分野を決定するための辞書管理情報と、前記ユーザ辞書を前記サーバへ送信する送信工程のプログラムコードと、

認識分野別に用意された複数種類の認識辞書群を前記サーバで保持する認識辞書保持工程のプログラムコードと、

前記クライアントから受信した辞書管理情報に対応する認識辞書を前記複数種類の認識辞書群から決定する決定工程のプログラムコードと、

前記決定工程で決定された認識辞書を少なくとも使用して、前記音声データの認識を行う認識工程のプログラムコードと

を備えることを特徴とするコンピュータ可読メモリ。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、クライアントで入力された音声をサーバで認識するクライアント・サーバ型の音声認識システム及びその制御方法、コンピュータ可読メモリに関するものである。

【 0 0 0 2 】

【従来の技術】

近年、キーボードやマウス等に加えて、音声が入力インタフェースとして用いられるようになってきた。

【 0 0 0 3 】

しかし、入力された音声を認識する音声認識では、音声認識の対象となる認識語彙が多くなる程、認識率が低下し、また、処理時間を必要とする。そのため、音声認識の対象となる認識語彙（例えば、読みと表記）を登録した認識辞書を複数持ち、入力対象や状況に応じて、認識辞書を切り換える（複数の認識辞書を同時に使用する場合もある）方法がとられている。

【 0 0 0 4 】

また、登録されていない語彙は認識することができないといった問題がある。この問題を解決する方法の一つに、ユーザ辞書（ユーザが音声認識の対象となる認識語彙を登録したもの）を用いる方法がある。

【 0 0 0 5 】

一方、リソースの小さい端末上で、音声認識を実現するために、クライアント



・サーバ型の音声認識システムが研究されている。

【0006】

【発明が解決しようとする課題】

上記の3つの技術は、既知であるが、これらの3つの技術を組み合わせたシステムは、現在のところ実現されていない。

【0007】

本発明は上記の課題を解決するためになされたものであり、クライアント・サーバ型の音声認識システムにおいて、ユーザからの要求に応じて、ユーザ辞書を使用することにより、音声入力の効率を向上し、システム全体の処理負荷を低減することができる音声認識システム及びその制御方法、コンピュータ可読メモリを提供することを目的とする。

【0008】

【課題を解決するための手段】

上記の目的を達成するための本発明による音声認識システムは以下の構成を備える。即ち、

クライアントで入力された音声をサーバで認識するクライアント・サーバ型の音声認識システムであって、

前記クライアントは、

音声を入力する音声入力手段と、

ユーザによって指定された認識対象語彙を登録することによって構成されるユーザ辞書を保持するユーザ辞書保持手段と、

前記音声入力手段で入力された音声データと、前記音声データの認識に使用する認識辞書の認識分野を決定するための辞書管理情報と、前記ユーザ辞書を前記サーバへ送信する送信手段と、

前記サーバは、

認識分野別に用意された複数種類の認識辞書群を保持する認識辞書保持手段と、

前記クライアントから受信した辞書管理情報に対応する認識辞書を前記複数種類の認識辞書群から決定する決定手段と、

前記決定手段で決定された認識辞書を少なくとも使用して、前記音声データの認識を行う認識手段とを備える。

【0009】

また、好ましくは、前記認識手段は、前記決定手段で決定された認識辞書と前記クライアントから受信したユーザ辞書を使用して、前記音声データを認識する。

【0010】

また、好ましくは、前記音声入力手段は、音声入力対象の入力フォームを表示する表示手段とを備え、

前記辞書管理情報は、前記入力フォームの種類を示す入力フォーム識別子である。

【0011】

また、好ましくは、前記辞書管理情報は、更に、前記音声データの認識において前記ユーザ辞書の使用の有無を示す情報を含む。

また、好ましくは、前記ユーザ辞書における認識対象語彙は、認識対象語彙の読みと表記とが対応づけられて構成される。

【0012】

また、好ましくは、前記ユーザ辞書は、更に、前記入力フォーム識別子の少なくとも1つと、前記認識対象語彙の読みと表記とが対応づけられて構成される。

【0013】

また、好ましくは、前記ユーザ辞書は、更に、前記複数種類の認識辞書群のそれぞれの認識分野を示す認識辞書識別子の少なくとも1つと、前記認識対象語彙の読みと表記とが対応づけられて構成される。

【0014】

上記の目的を達成するための本発明による音声認識システムの制御方法は以下の構成を備える。即ち、

クライアントで入力された音声サーバで認識するクライアント・サーバ型の音声認識システムの制御方法であって、

音声を入力する音声入力工程と、

ユーザによって指定された認識対象語彙を登録することによって構成されるユーザ辞書を前記クライアントで保持するユーザ辞書保持工程と、

前記音声入力工程で入力された音声データと、前記音声データの認識に使用する認識辞書の認識分野を決定するための辞書管理情報と、前記ユーザ辞書を前記サーバへ送信する送信工程と、

認識分野別に用意された複数種類の認識辞書群を前記サーバで保持する認識辞書保持工程と、

前記クライアントから受信した辞書管理情報に対応する認識辞書を前記複数種類の認識辞書群から決定する決定工程と、

前記決定工程で決定された認識辞書を少なくとも使用して、前記音声データの認識を行う認識工程と、

を備える。

【0015】

上記の目的を達成するための本発明によるコンピュータ可読メモリは以下の構成を備える。即ち、

クライアントで入力された音声をサーバで認識するクライアント・サーバ型の音声認識システムの制御のプログラムコードが格納されたコンピュータ可読メモリであって、

音声を入力する音声入力工程と、

ユーザによって指定された認識対象語彙を登録することによって構成されるユーザ辞書を前記クライアントで保持するユーザ辞書保持工程のプログラムコードと、

前記音声入力工程で入力された音声データと、前記音声データの認識に使用する認識辞書の認識分野を決定するための辞書管理情報と、前記ユーザ辞書を前記サーバへ送信する送信工程のプログラムコードと、

認識分野別に用意された複数種類の認識辞書群を前記サーバで保持する認識辞書保持工程のプログラムコードと、

前記クライアントから受信した辞書管理情報に対応する認識辞書を前記複数種

類の認識辞書群から決定する決定工程のプログラムコードと、

前記決定工程で決定された認識辞書を少なくとも使用して、前記音声データの認識を行う認識工程のプログラムコードと

を備える。

【 0 0 1 6 】

【発明の実施の形態】

以下、図面を参照して本発明の好適な実施形態を詳細に説明する。

【実施形態 1】

図 1 は実施形態 1 の音声認識システムのハードウェア構成を示す図である。

【 0 0 1 7 】

CPU 1 0 1 は、クライアント 1 0 0 全体を統括制御するものであり、ROM 1 0 2 に格納されているプログラムを RAM 1 0 3 に読み出し、その読み出したプログラムに基づいて、各種処理動作を実行する。ROM 1 0 2 は、CPU 1 0 1 で実行する処理の各種プログラムを格納している。RAM 1 0 3 は、ROM 1 0 2 に格納されている各種プログラムの実行に必要な記憶領域を提供する。

【 0 0 1 8 】

二次記憶装置 1 0 4 は、OS や各種プログラムを格納している。但し、クライアント 1 0 0 をパーソナルコンピュータ等の汎用装置ではなく、専用装置で構成する場合には、ROM 1 0 2 内に OS や各種プログラムを格納しても構わない。この格納されたプログラムを RAM 1 0 3 に読み出すことによって、CPU 1 0 1 が処理を実行することが可能である。また、二次記憶装置 1 0 4 としては、ハードディスク装置、フロッピーディスクドライブ、CD-ROM 等がある。つまり、記憶媒体は如何なるものであっても構わない。

【 0 0 1 9 】

ネットワーク I / F ( インタフェース ) 1 0 5 は、サーバ 2 0 0 のネットワーク I / F 2 0 5 と接続される。

【 0 0 2 0 】

入力装置 1 0 6 は、マウスやキーボード、マイク等で構成され、CPU 1 0 1 で実行される処理の各種指示の入力を可能とし、これら複数の機器を同時に接続

して使用できる。出力装置107は、ディスプレイ（CRTやLCD等）で構成され、入力装置106より入力された情報や、CPU101で実行される各種処理で制御される表示画面を表示する。バス108は、クライアント100の各種構成要素を相互に接続する。

#### 【0021】

CPU201は、サーバ200全体を統括制御するものであり、ROM202に格納されているプログラムをRAM203に読み出し、その読み出したプログラムに基づいて、各種処理動作を実行する。ROM202は、CPU201で実行する処理の各種プログラムを格納している。RAM203は、ROM202に格納されている各種プログラムの実行に必要な記憶領域を提供する。

#### 【0022】

二次記憶装置204は、OSや各種プログラムを格納している。但し、サーバ200が、パーソナルコンピュータ等の汎用装置ではなく、専用装置で構成する場合には、ROM202内にOSや各種プログラムを格納しても構わない。この格納されたプログラムをRAM203に読み出すことによってCPU201が処理を実行することが可能である。また、二次記憶装置204としては、ハードディスク装置、フロッピーディスクドライブ、CD-ROM等がある。つまり、記憶媒体は如何なるものであっても構わない。

#### 【0023】

ネットワークI/F205は、クライアント100のネットワークI/F105と接続される。バス106は、サーバ200の各種構成要素を相互に接続する。

#### 【0024】

次に、実施形態1の音声認識システムの機能構成について、図2を用いて説明する。

#### 【0025】

図2は実施形態1の音声認識システムの機能構成を示すブロック図である。

#### 【0026】

音声入力部121は、マイク（入力装置106）からユーザが発生した音声を

入力し、その入力された音声認識対象の音声データ（音声認識用データ）のA/D変換を行う。通信部122は、ユーザ辞書124a、音声認識用データ124b、辞書管理情報124c等をサーバ200に送信し、送信した音声認識用データ124bに対する音声認識結果等をサーバ200から受信する。

## 【0027】

通信部221は、ユーザ辞書124a、音声認識用データ124b、辞書管理情報124c等をクライアント100から受信し、受信した音声認識用データ124bに対する音声認識結果等をクライアント100に送信する。

## 【0028】

表示部123は、サーバ200から受信した音声認識結果を、例えば、出力装置108で表示される本音声認識システムで実行される処理によって表示される表示画面上の入力フォーム等に格納して表示する。

## 【0029】

辞書管理部223は、入力対象や状況等に応じて、音声認識に使用する認識分野別（例えば、名前用、住所用、英数字記号用等）に用意された複数種類の認識辞書群225（認識辞書1～認識辞書N、N：正の整数）及びクライアント100から受信したユーザ辞書124a）を切り換えて、選択する（複数種類の辞書を同時に使用する場合もある）。

## 【0030】

尚、この複数種類の認識辞書群225は、後述するクライアント100から送信されてくる辞書管理情報124c（入力フォーム識別子）毎に用意されており、各認識辞書225にはその認識辞書の認識分野を示す認識辞書識別子が付与されており、辞書管理部223は、この認識辞書識別子と入力フォーム識別子を対応づけた、図5に示すような識別子テーブル223aを管理している。

## 【0031】

音声認識部224は、クライアント100から受信した音声認識用データ124b、辞書管理情報124cに基づいて辞書管理部223が音声認識用に指定する認識辞書225、ユーザ辞書124aを使用して音声認識を行う。

## 【0032】

尚、ユーザ辞書 1 2 4 a は、ユーザが音声認識の対象となる認識語彙を登録したものであり、実施形態 1 の場合、例えば、図 3 に示すように、認識対象語彙の読みと表記を対応づけて構成される。

【 0 0 3 3 】

また、音声認識用データ 1 2 4 b は、音声入力部 1 2 1 において A/D 変換された音声データであっても良いし、その音声データを符号化したデータであっても良い。

【 0 0 3 4 】

また、辞書管理情報 1 2 4 c は、入力対象等を示す情報である。この辞書管理情報 1 2 4 c は、例えば、図 4 に示す実施形態 1 の音声認識システムで表示される音声入力用画面を構成する各入力フォームに、入力された音声をサーバ 2 0 0 が音声認識し、その音声認識結果に対応するテキストデータを入力する場合において、図 5 に示すような、入力フォームの種類を示す識別子（入力フォーム識別子）である。そして、クライアント 1 0 0 は、この入力フォーム識別子を辞書管理情報 1 2 4 c としてサーバ 2 0 0 に送信し、サーバ 2 0 0 では、辞書管理部 2 2 3 において、識別子テーブル 2 2 3 a を参照して、受信した入力フォーム識別子に対応する認識辞書識別子を取得し、音声認識に使用する認識辞書 2 2 5 を決定する。

【 0 0 3 5 】

次に、実施形態 1 の音声認識システムで実行される処理について、図 6 を用いて説明する。

【 0 0 3 6 】

図 6 は実施形態 1 の音声認識システムで実行される処理を示すフローチャートである。

【 0 0 3 7 】

ステップ S 1 0 1 では、クライアント 1 0 0 は、ユーザ辞書 1 2 4 a をサーバ 2 0 0 に送信する。

【 0 0 3 8 】

ステップ S 2 0 1 では、サーバ 2 0 0 は、ユーザ辞書 1 2 4 a をクライアント

1 0 0 から受信する。

【0 0 3 9】

ステップ S 1 0 2 では、クライアント 1 0 0 は、音声入力対象の入力フォームに音声入力が行われた場合に、その入力フォームの入力フォーム識別子を辞書管理情報 1 2 4 c として、サーバ 2 0 0 に送信する。

【0 0 4 0】

ステップ S 2 0 2 では、サーバ 2 0 0 は、辞書管理情報 1 2 4 c として、入力フォーム識別子をクライアント 1 0 0 から受信する。

【0 0 4 1】

ステップ S 2 0 3 では、サーバ 2 0 0 は、辞書管理情報 1 2 4 c で識別テーブル 2 2 3 a を参照して、受信した入力フォーム識別子に対応する認識辞書識別子を取得し、音声認識に使用する認識辞書 2 2 5 を決定する。

【0 0 4 2】

ステップ S 1 0 3 では、クライアント 1 0 0 は、各入力フォームに入力するテキストデータとして音声入力された音声認識用データ 1 2 4 b をサーバ 2 0 0 に送信する。

【0 0 4 3】

ステップ S 2 0 4 では、サーバ 2 0 0 は、各入力フォームに対応する音声認識用データ 1 2 4 b をクライアント 1 0 0 から受信する。

【0 0 4 4】

ステップ S 2 0 5 では、サーバ 2 0 0 は、音声認識用データ 1 2 4 b に対し、辞書管理部 2 2 3 で音声認識用に指定された認識辞書 2 2 5、ユーザ辞書 1 2 4 を用いて、音声認識部 2 2 4 で音声認識を行う。

【0 0 4 5】

尚、実施形態 1 では、音声認識部 2 2 4 による音声認識の際には必ず、クライアント 1 0 0 からサーバ 2 0 0 に送信されたユーザ辞書 1 2 4 a 中の認識語彙を全て、認識対象語彙としている。

【0 0 4 6】

ステップ S 2 0 6 では、サーバ 2 0 0 は、音声認識部 2 2 4 から得られる音声



認識結果をクライアント 1 0 0 に送信する。

【 0 0 4 7 】

ステップ S 1 0 4 では、クライアント 1 0 0 は、各入力フォームに対する音声認識結果をサーバ 2 0 0 から受信し、それぞれ対応する入力フォームに音声認識結果に対応するテキストデータを格納する。

【 0 0 4 8 】

ステップ S 1 0 5 では、クライアント 1 0 0 は、処理を終了するか否かを判定する。処理を終了しない場合（ステップ S 1 0 5 で N O）、ステップ S 1 0 2 に戻り、処理を繰り返す。一方、処理を終了する場合（ステップ S 1 0 5 で Y E S）、サーバ 2 0 0 に処理を終了する旨を通知し、処理を終了する。

【 0 0 4 9 】

ステップ S 2 0 7 では、サーバ 2 0 0 は、クライアント 1 0 0 から処理の終了の指示があるか否かを判定する。処理の指示がない場合（ステップ S 2 0 7 で N O）、ステップ S 2 0 2 に戻り、処理を繰り返す。一方、処理の終了の指示がある場合（ステップ S 2 0 7 で Y E S）、処理を終了する。

【 0 0 5 0 】

尚、上記処理では、音声入力対象の入力フォームに対する音声入力が行われた時に、その入力フォームに対応する辞書管理情報 1 2 4 c をクライアント 1 0 0 からサーバ 2 0 0 へ送信する構成としたが、音声入力対象の入力フォームが、入力装置 1 0 6 の指示によってフォーカス（音声入力対象の入力フォームが確定）された時に、送信するように構成しても良い。

【 0 0 5 1 】

また、サーバ 2 0 0 では、音声認識用データ 1 2 4 b を全て受信してから音声認識を行う構成としたが、ある入力フォームに入力するテキストデータとして音声入力が行われる毎に、その音声認識用データ 1 2 4 b を逐次サーバ 2 0 0 に送信し、リアルタイムで音声認識を行うように構成しても良い。

【 0 0 5 2 】

以上説明したように、実施形態 1 によれば、クライアント・サーバ型の音声認識システムにおいて、音声認識用データ 1 2 4 b の音声認識に用いる適切な認識

辞書 2 2 5 とユーザ辞書 1 2 4 a を併用して、サーバ 2 0 0 において音声認識を実行することで、クライアント 1 0 0 の音声認識に関わる処理負荷、記憶資源の使用を低減し、かつサーバ 2 0 0 における音声認識精度を向上することができる。

#### 〔実施形態 2〕

尚、実施形態 1 において、ユーザ辞書 1 2 4 a 中の認識語彙が発生しない場合には、ユーザ辞書 1 2 4 a を使用する必要がないため、サーバ 2 0 0 は、クライアント 1 0 0 からユーザ辞書 1 2 4 a を使用するという要求がある場合にのみ、ユーザ辞書 1 2 4 a 中の認識語彙を全て認識対象語彙としても良い。

#### 【0 0 5 3】

この場合は、例えば、辞書管理情報 1 2 4 c として、ユーザ辞書 1 2 4 a の使用の有無を示すフラグを追加することで、ユーザ辞書 1 2 4 a の使用の有無をサーバ 2 0 0 に通知する。

#### 〔実施形態 3〕

また、入力対象や状況等によって、ユーザ辞書 1 2 4 a 中の認識対象語彙中に使用しない語彙があるため、入力対象や状況に応じて、ユーザ辞書 1 2 4 a 中の特定の認識語彙のみを認識対象語彙としても良い。

#### 【0 0 5 4】

この場合は、例えば、図 7 に示すように、認識語彙ごとに、入力フォーム識別子を指定して管理することにより、音声入力された入力フォームの入力フォーム識別子を持つ認識語彙のみを、認識対象語彙とすることができる。また、ある認識語彙に対して複数の入力フォーム識別子を指定しても良い。加えて、図 8 に示すように、入力フォーム識別子の代わりに、認識辞書識別子を指定して管理しても良い。

#### 〔実施形態 4〕

実施形態 2 と実施形態 3 を組み合わせることにより、音声認識部 4 による音声認識処理の更なる効率化を図ることができる。

#### 〔実施形態 5〕

本装置は、上記の処理のほとんどがプログラムによって実現できる。先に説明

したように、装置としては、パーソナルコンピュータ等の汎用装置で良いわけであるから、本発明は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体をシステムあるいは装置に供給し、そのシステムあるいは装置のコンピュータが記憶媒体に格納されたプログラムコードを読み出し実行することによっても実現はできるのはもちろんである。この場合、記憶媒体から読み出されたプログラムコード自体が、前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。プログラムコードを供給するための記憶媒体としては、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、磁気テープ、不揮発性のメモリカード、ROM等を用いることができる。

## 【 0 0 5 5 】

また、本発明は、前述したプログラムコードを記録した記録媒体をコンピュータに供給し、そのコンピュータ上で稼動しているOS等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現されていることによっても達成できる。さらに、記憶媒体から読み出されたプログラムコードが、コンピュータによって挿入された機能拡張ボードや機能拡張ユニットに備わるメモリに書き込まれた後、機能拡張ボードや機能拡張ユニットに備わるCPU等がプログラムコードの指示に基づいて実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される。本発明を上記記憶媒体に適用する場合、その記憶媒体には前述の図3のフローチャートに対応するプログラムコードを格納することになる。

## 【 0 0 5 6 】

## 【発明の効果】

以上説明したように、本発明によれば、クライアント・サーバ型の音声認識システムにおいて、ユーザからの要求に応じて、ユーザ辞書を使用することにより、音声入力効率を向上し、システム全体の処理負荷を低減することができる音声認識システム及びその制御方法、コンピュータ可読メモリを提供できる。

## 【図面の簡単な説明】

## 【図1】

実施形態 1 の音声認識システムのハードウェア構成を示す図である。

【図 2】

実施形態 1 の音声認識システムの機能構成を示すブロック図である。

【図 3】

実施形態 1 のユーザ辞書の構成を示す図である。

【図 4】

実施形態 1 の音声入力用画面を示す図である。

【図 5】

実施形態 1 の識別子テーブルを示す図である。

【図 6】

実施形態 1 の音声認識システムで実行される処理を示すフローチャートである。

【図 7】

実施形態 3 の入力フォーム識別子を付加したユーザ辞書の構成を示す図である。

【図 8】

実施形態 3 の認識辞書識別子を付加したユーザ辞書の構成を示す図である。

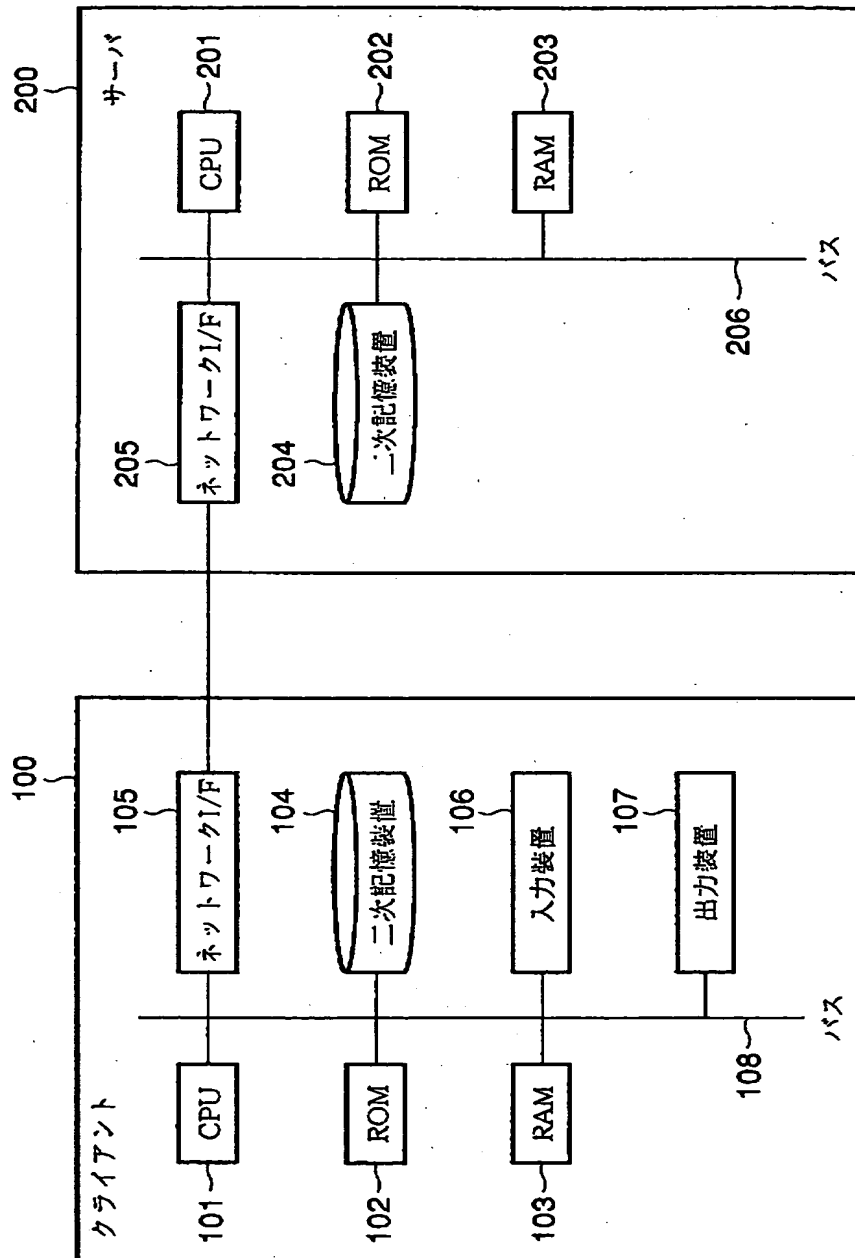
【符号の説明】

- 100 クライアント
- 101 CPU
- 102 ROM
- 103 RAM
- 104 二次記憶装置
- 105 ネットワーク I/F
- 106 入力装置
- 107 出力装置
- 108 バス
- 121 音声入力部
- 122 通信部

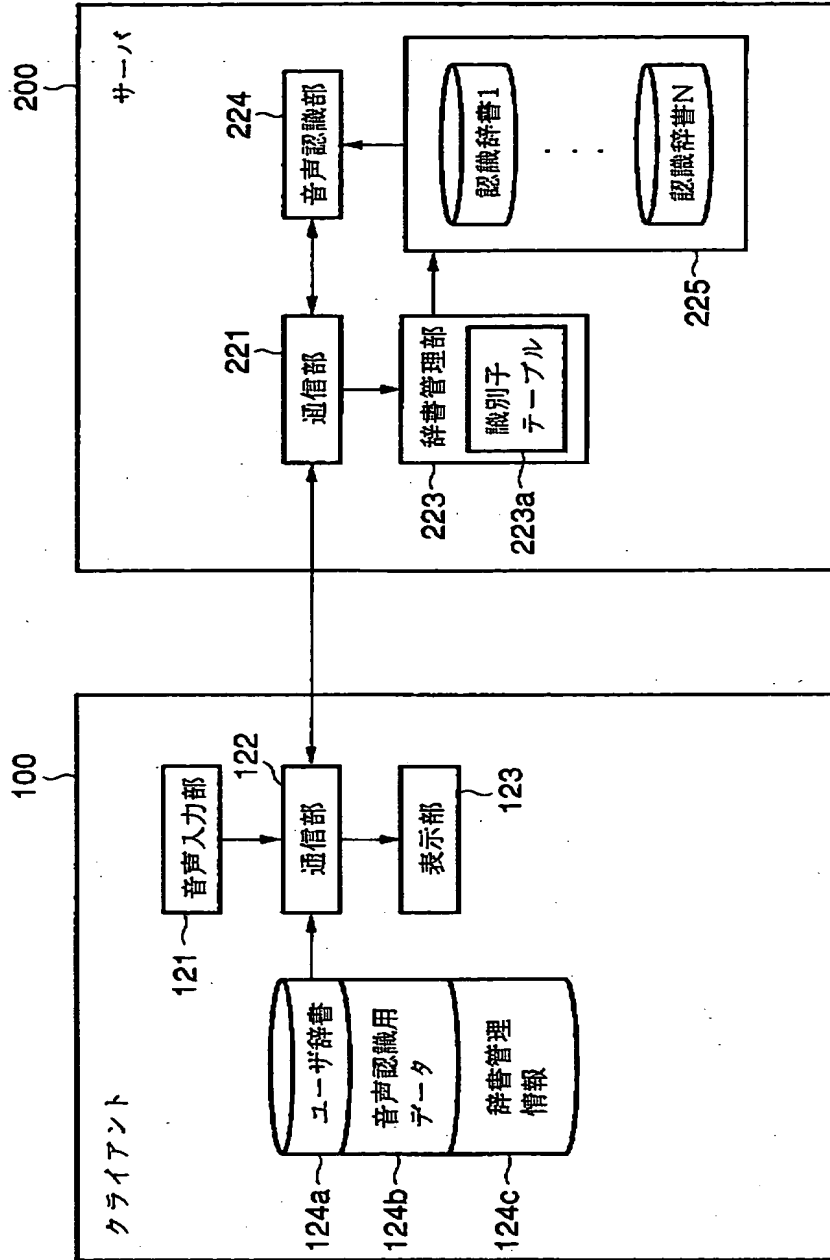
- 123 表示部
- 124 a ユーザ辞書
- 124 b 音声認識用データ
- 124 c 辞書管理情報
- 200 サーバ
- 201 CPU
- 202 ROM
- 203 RAM
- 204 二次記憶装置
- 205 ネットワーク I/F
- 206 バス
- 221 通信部
- 223 辞書管理部
- 223 a 識別子テーブル
- 224 音声認識部
- 225 認識辞書

【書類名】 図面

【図 1】



【図 2】



【図 3】

ユーザ辞書

読み	表記
スズキタロー	鈴木太郎
ジタクジュウショ	神奈川県川崎市中原区今井上町999
ジタクデンワ ケイタイデンワ	045-222-2222 090-1111-1111
ファックス	045-333-3333
メール	suzuki@xxx.yyy.co.jp



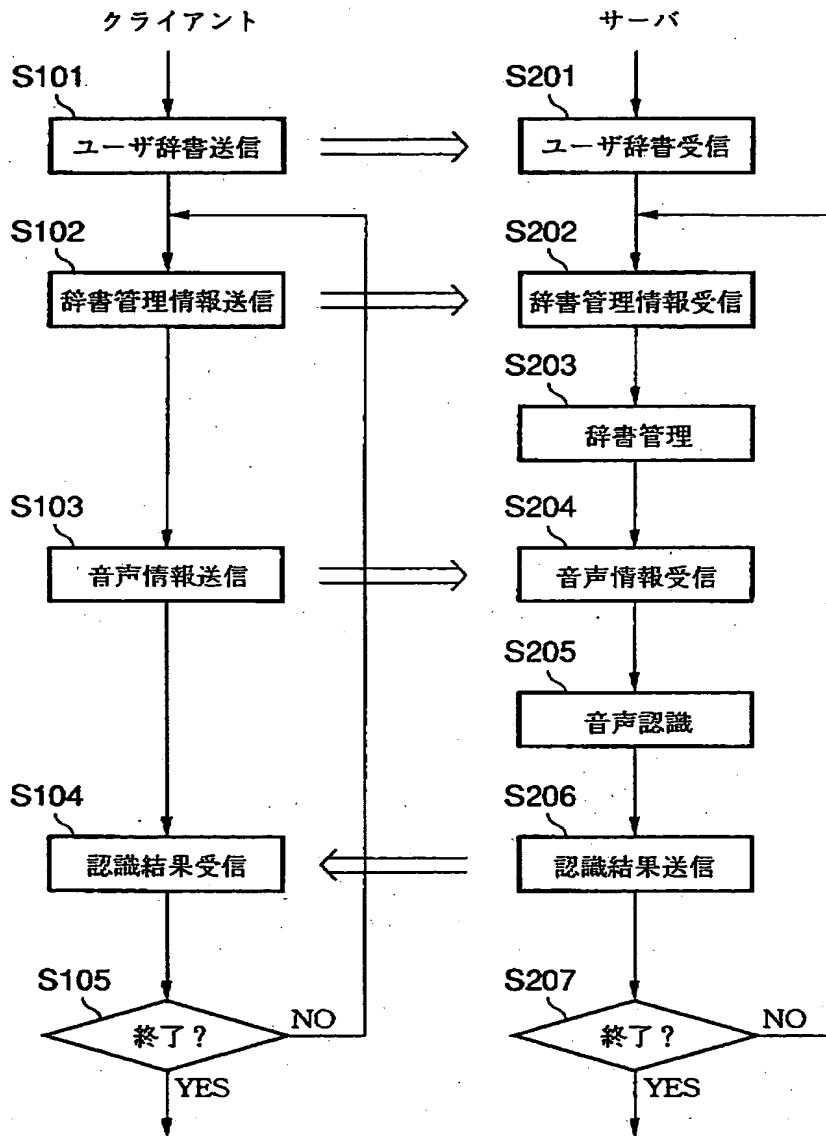
【図4】

氏名	<input type="text"/>
住所	<input type="text"/>
TEL	<input type="text"/>
FAX	<input type="text"/>
E-mail	<input type="text"/>
その他	<input type="text"/>

【図 5】

入力フォーム識別子	認識辞書識別子
氏名	名前
住所	住所
TEL	英数字記号
FAX	英数字記号
E-mail	英数字記号
その他	ディクテーション

【図 6】



【図 7】

ユーザ辞書

読み	表記	入力フォーム識別子
スズキタロー	鈴木太郎	氏名
ジタクジユウショ	神奈川県川崎市中原区今井上町999	住所
ジタクデデンワ ケイタイデンワ	045-222-2222 090-1111-1111	TEL TEL
ファックス	045-333-3333	FAX
メール	suzuki@xxx.yyy.co.jp	E-mail
スズキタロー	鈴木太郎	その他

【図8】

ユーザ辞書

読み	表記	認識辞書識別子
スズキタロー	鈴木太郎	名前
ジタクジュウシヨ	神奈川県川崎市中原区今井上町53	住所
ジタクデンワ ケイタイデンワ	045-222-2222 090-1111-1111	英数字記号 英数字記号
フックス	045-333-3333	英数字記号
メール	suzuki@xxx.yyy.co.jp	英数字記号
スズキタロー	鈴木太郎	ディクテーション

【書類名】 要約書

【要約】

【課題】 ユーザからの要求に応じて、ユーザ辞書を使用することにより、音声入力の効率を向上し、システム全体の処理負荷を低減することができる音声認識システムを提供する。

【解決手段】 ユーザによって指定された認識対象語彙の読みと表記を対応づけて構成されるユーザ辞書 1 2 4 a と、入力された音声認識用データ 1 2 4 b と、音声認識用データ 1 2 4 b の認識に使用する認識辞書の認識分野を決定するための辞書管理情報 1 2 4 c をサーバ 2 0 0 へ通信部 1 2 2 を介して送信する。サーバ 2 0 0 は、辞書管理部 2 2 3 によって識別子テーブル 2 2 3 a を参照し、クライアント 1 0 0 から受信した辞書管理情報に対応する認識辞書を複数種類の認識辞書群 2 2 5 から決定する。音声認識部 2 2 4 で、決定された認識辞書を少なくとも使用して、音声認識用データ 2 2 4 b の認識を行う。そして、その認識結果をクライアント 1 0 0 へ通信部 2 2 1 を介して送信する。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号

[000001007]

1. 変更年月日 1990年 8月30日

[変更理由] 新規登録

住 所 東京都大田区下丸子3丁目30番2号

氏 名 キヤノン株式会社